GROW Working Group
Internet-Draft
Intended status: Informational
Expires: January 7, 2010

X. Xu
Huawei
P. Francis
MPI-SWS
R. Raszuk
Cisco Systems

July 06, 2009

GRE and IP-in-IP Tunnels for Virtual Aggregation
draft-ietf-grow-va-gre-00.txt

Status of this Memo

Copyright Notice

Abstract

The document "FIB Suppression with Virtual Aggregation" [I-D.grow-va] describes how FIB size may be reduced. That draft refers generically to tunnels, and leaves it to other documents to define the tunnel establishment methods for specific tunnel types. This document provides those definitions for GRE and IP-in-IP tunnels.

Table of Contents

1. Introduction

This document specifies how to signal and use GRE and IP-in-IP tunnels as required by [I-D.grow-va], "FIB Suppression with Virtual Aggregation". This document adopts the terminology of [I-D.grow-va]. This document covers the behavior for both VA routers and legacy routers.

1.1. Requirements notation

The key words "must", "must NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in [RFC2119].

2. Tunneling Requirements

According to [I-D.grow-va], VA has the following tunnel-related requirements. The requirement numbers here (R1 - R5) are cited by [I-D.grow-va].

R1: Legacy routers and APRs must be able to detunnel packets addressed to themselves at their BGP NEXT_HOP address. They must be able to convey the tunnel parameters needed by other routers to initiate these tunneled packets.

R2: Border VA routers must be able to detunnel packets targeted to neighboring remote ASBRs. They must be able to forward these packets to the targeted remote ASBR without doing a FIB lookup. They must be able to convey the tunnel parameters needed by other routers to initiate these tunneled packets.

R3: VA routers must be able to initiate tunneled packets targeted to any BGP NEXT_HOP address (i.e. those for APRs, legacy routers, or remote ASBRs).

R4: Legacy routers may optionally be able to initiate tunneled packets targeted to any BGP NEXT_HOP address (i.e. those for APRs, legacy routers, or remote ASBRs). The GRE and IP-in-IP tunnels

defined in this document do not have this capability.
R5: All routers must be able to forward all tunneled packets.

3. Tunneling Specification for GRE and IP-in-IP

This document distinguishes between the terms "tunnel endpoint", and "tunnel target". The tunnel endpoint is the router that detunnels the packet (i.e. strips out the outer header and forwards the no-longer-tunneled packet). The tunnel target, on the other hand, is the router to which the packet is going. This distinction manifests itself in the case of requirement R2. That is, a local ASBR (border

router) is a VA router, and it detunnels packets. The remote ASBR, however, is the router to which the packet is ultimately targeted. Here, the tunnel endpoint is the local ASBR, and the tunnel target is the remote ASBR.

The IP address of the outer header for GRE and IP-in-IP tunnels is always addressed to the tunnel endpoint. If the tunnel endpoint and the tunnel target are the same router (as with the case in requirement R1), then the tunnel type may be GRE or IP-in-IP. If the former, then the Key field may or may not be used.

If the tunnel endpoint and the tunnel target are different routers (as is the case in requirement R2), then this document specifies two tunneling approaches. One requires the use of GRE, where the Key field is used to identify the tunnel target to the tunnel endpoint. This is called Key-based identification. The other does not require the use of the Key field, and therefore can be either GRE or IP-in-IP. Instead of using the Key field to identify the tunnel target, a distinct destination IP address is used per tunnel target (remote ASBR) to identify the tunnel target to the tunnel endpoint. This is called address-based identification.

The following examples clarify these two cases. Assume a local ASBR has two remote ASBR neighbors, with addresses 2.2.2.2 and 3.3.3.3 respectively.

In the case of Key-based identification, the local ASBR would assign two GRE Key values, one for each remote ASBR neighbor. The local ASBR would advertise it's own IP address (say 10.1.1.1) as the BGP

NEXT_HOP. All GRE packets would arrive at 10.1.1.1 (the tunnel endpoint), which would then look at the Key value to determine whether to forward the packet to 2.2.2.2 or 3.3.3.3. Note that no FIB lookup is necessary.

In the case of Address-based identification, the local ASBR would be reachable at a block of IP addresses, say 10.1.1/24. The local ASBR would assign one address from the block for each neighbor remote ASBR. For instance, it could assign the address 10.1.1.2 to remote ASBR 2.2.2.2, and assign the address 10.1.1.3 to remote ASBR 3.3.3.3. Likewise, when advertising NLRI reachable through 2.2.2.2, it would advertise a BGP NEXT_HOP of 10.1.1.2. Packets received at the tunnel endpoint 10.1.1.2 would be forwarded to 2.2.2.2 without a FIB lookup. When advertising NLRI reachable through 3.3.3.3, it would advertise a BGP NEXT_HOP of 10.1.1.3. Packets received at the tunnel endpoint 10.1.1.3 would be forwarded to 3.3.3.3 without a FIB lookup.

3.1. Conveying GRE and IP-in-IP tunnel parameters

This document uses two BGP attributes defined in [RFC5512] to convey the parameters necessary for routers to initiate tunneled packets (i.e. requirement R3). The first attribute, the BGP Encapsulation Extended Community (BGPencap-Attribute), is used when the tunnel type is IP-in-IP or GRE without the Key field. The second BGP attribute, the Tunnel Encapsulation Attribute (TEncap-Attribute), is used when the tunnel type is GRE with a Key field. In either case, routers must tunnel packets to the NEXT_HOP address in the BGP update.

3.1.1. Usage of the RFC5512 Attributes

Legacy routers are defined here as routers that do not do FIB suppression, but do implement RFC5512. Legacy routers must be configured to attach the BGPencap-Attribute to all iBGP updates, and to detunnel packets addressed to the NEXT_HOP address advertised by the legacy router. This satisfies requirement R1 for legacy routers.

In the case where VA routers used Key-based identification, the BGP NEXT_HOP must be set to the local ASBR, GRE must be used, and the TEncap-Attribute must be included. The GRE Key field must be set to

a value unique for the remote ASBR to which the packet must be delivered. If the Key value for a given remote ASBR is modified, then both the old and new Key values must identify the remote ASBR in received packets until the new iBGP updates are fully disseminated. This satisfies requirement R2.

In the case where VA routers use Address-based identification, the router must have a distinct locally assigned address for each neighbor remote ASBR. The BGP NEXT_HOP field is set to this locally assigned address. This also satisfies requirement R2.

If the VA router is an APR, then for tunnels associated with the VP route, where the BGP NEXT_HOP is that of the VA router itself, GRE may or may not be used. If it is used, then the APR must have a way to distinguish tunnels targeted at itself from tunnels targeted to a neighbor remote ASBR. Where Key-based identification is used, this can be done by assigning a unique Key value (i.e. one not assigned to a remote ASBR). Where address-based identification is used, this can be done by using a local IP address not assigned to a remote ASBR. This satisfies requirement R1 for VA routers.

All VA routers must use the tunnels described in the tunnel attributes to forward packets to resolved BGP NEXT_HOPs (requirement R3).

4. IANA Considerations

There are no IANA considerations.

5. Security Considerations

There are no new security considerations beyond those already described in [I-D.grow-va].

6. Normative References

[I-D.grow-va]
Francis, P., Xu, X., Ballani, H., Jen, D., Raszuk, R., and

L. Zhang, "FIB Suppression with Virtual Aggregation", draft-ietf-grow-va-00 (work in progress), May 2009.

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

[RFC5512] Mohapatra, P. and E. Rosen, "BGP Encapsulation SAFI and BGP Tunnel Encapsulation Attribute", RFC 5512, April 2009.

Authors' Addresses

Xiaohu Xu
Huawei Technologies
No.3 Xinxi Rd., Shang-Di Information Industry Base, Hai-Dian District
Beijing, Beijing 100085
P.R.China

Phone: +86 10 82836073
Email: xuxh@huawei.com

Paul Francis
Max Planck Institute for Software Systems
Gottlieb-Daimler-Strasse
Kaiserslautern 67633
Germany

Phone: +49 631 930 39600
Email: francis@mpi-sws.org

Robert Raszuk
Cisco Systems
170 West Tasman Drive
San Jose, CA 95134
US

Email: raszuk@cisco.com